Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, No. 5, July : 2023

COMPARISON OF RECENT MEDICAL IMAGE RETRIEVAL METHODS

N.Rajesh Kumar, PhD Scholar, ECE Department, University College of Engineering, Osmania University, Hyderabad, TS, INDIA
Dr. Y. Ravi Kumar, Scientist G Retd, DLRL, TS, INDIA
raj.naskanti@gmail.com ; dr.ravikumaryeda@gmail.com

Abstract: *The capacity of a retrieval system for the purpose of obtaining attributes through the utilization of that system's feature descriptor is the key criterion that is used to measure the efficiency of that system. The process of diagnosis based on medical images is laborious, and it is possible for medical professionals to miss minor lesions in a variety of medical images because of short attention span of the human visual system. This can have a negative impact on the medical therapy that is provided. The conventional text- and metadata-based methods of picture retrieval are insufficient for dealing with automatically accumulated image collections and personal digital libraries. In most cases, these methods do not include thorough descriptions that may be utilized for the purpose of searching for the necessary image. In order to provide a more fruitful image search within digital libraries, content-based image retrieval (CBIR) techniques are necessary. This paper concentrates on the comparison of various medical image retrieval methods and their efficiencies.*
Keywords: *Content Based Image Retrieval, Medical Image, Neural Network, quality parameters*.

I. Introduction

Over the course of the last several decades, computer-aided assessment methods and techniques have gained widespread use as a means of improving clinical therapy [1, 2]. These cutting-edge tools provide assistance to medical professionals in a variety of practice areas, including clinical therapy for any particular illness or injury. These various medical imaging technologies allow for better diagnosis and treatment by providing a visual glimpse into distinct human parts that are normally concealed from view. Because of the intricate structure of various body parts, analyzing medical images is a difficult undertaking, and effective interpretation requires the assistance of trained medical professionals [3-6]. The CBIR system does not rely on the description being provided in the form of text. The concepts of machine learning confer learning capabilities on a computer in the absence of any explicit programming. The primary focus of machine learning is on the modifications that take place after the system is presented with new data. Machine learning and the deployment of iterative approaches are the ones that are seen most frequently in CBIR.

During the course of ordinary clinical procedures, hospitals and other medical institutes all around the globe generate and maintain a wide variety of datasets of biomedical images (BMI).These datasets include the images that were generated by a variety of biomedical modalities such as X-Ray, ultrasound (US), computed tomography (CT), and magnetic resonance imaging (MRI).These many imaging modalities give doctors the ability to accurately diagnose their patients [7]. Pattern recognition and image processing frequently make use of a technique known as texture analysis. The wavelet transform makes it feasible to effectively separate and extract both the low-frequency approximation information as well as the high-frequency detail texture information from medical imagery. The latter can reveal the complexities of the medical image and get more important information that can be retrieved. Because the majority primitives that may be seen on the surface of damaged organs are uneven, it is possible to conduct a study of the texture using the Tamura method, which is a technique that is frequently utilized in statistical analysis. As a result of this, in order to create a new feature vector group we are going to merge the Tamura texture feature with the average coefficients that were generated through the wavelet decomposition [8]. In addition to that, the medical picture retrieval process will make use of the Hausdorff distance of matched point sets.

The use of digital photography is becoming increasingly widespread for a variety of reasons. Only a few examples are provided here, but they include private photo collections, medical imaging, and geographical information systems. The amount of digital image collections is quickly expanding as a result of a combination of factors, including the rise in computing power and the fall in the price of storage medium [9-12]. There is a need for ways that can present us with the information in a way that is both efficient and convenient, as well as strategies that enable us to access and retrieve the enormous quantity of information that is embedded in these collections. Even with private collections, straightforward manual browsing is becoming increasingly burdensome. The retrieval of images automatically is unavoidable.

The phrase "content-based image retrieval" refers to a collection of methods that are used to retrieve semantically meaningful images from an image database. These methods are based on automatically derived image features. CBIR's primary objective is to maximize efficiency during the image indexing and retrieval process, with the end goal of cutting down on the amount of human involvement that is required in the indexing procedure. The machine needs to be able to get photos from a database without relying on any assumptions made by humans about particular fields. One of the primary responsibilities of CBIR systems is the comparison of images for similarity, which entails deriving the feature signatures of each image based on the values of its pixels and providing the rules for comparing images. These attributes are what's known as the image representation structure, and they're used to determine how similar a picture is to others that have been stored in the database. Calculating the difference between an image's feature components and the corresponding feature components of another image is the method used to compare images [13-16]. The field of image processing and computer vision is where CBIR finds inspiration for many of its methodologies. The term "image processing" refers to a considerably broader industry that encompasses a variety of subfields such as image enhancement, compression, transmission, and interpretation.

II. Recent Medical retrieval methods

Texture Features Fusion (**TFF):** In this study, the eigenvector of the image retrieval process is formed by extracting the texture information of a medical image after it has been decomposed twice using wavelet analysis. A one-level wavelet is used to perform a decomposition on the medical image, which results in the production of four sub-images, each of which has a size that is one-fourth that of the original medical image. According to Figure 3, the low-frequency information of the image is located in the upper left corner LL1, while the high-frequency information of the image is located throughout the rest of the image, with LH1 representing the vertical component of the image, HL1 representing the horizontal component of the image, and HH1 representing the diagonal component of the image. After performing the 1-level wavelet decomposition on medical pictures, a second round of the 2-level wavelet decomposition is carried out on the low-frequency component LL1 of the images, together with the low frequency component LL2 and the six high-frequency components. In addition, a diagram depicting the 1-level and 2-level discrete wavelet decomposition of the brain MRI pictures and lung CT images is presented here [17].In order to construct a wavelet decomposition tree, the discrete wavelet transform performs a multilevel decomposition of low-frequency components of the primary medical image. Last but not least, the low-frequency approximation information and the high-frequency detail texture information of the medical imagery are successfully separated from one another and retrieved from one another.

In everyday life, there are typically 15 different types of wavelets that are employed. In the field of medical imaging feature extraction, there are 7 different discrete wavelet types that can be used. These are Haar, Daubechies, Bi-orthogonal, Coiflet, Symlets, Dmeyer, and Reverse Bior. An essential component of the procedure for the recovery of images is the extraction of features. Image retrieval also includes the evaluation of similarities between the images. At the moment, distance and correlation are the most prevalent methods that are utilized in order to measure similarity. When comparing two photographs, the degree to which their feature points are separated by a shorter distance indicates how closely they resemble one another. The Euclidean distance, the Hausdorff distance, the Manhattan distance, and the EMD distance are the four approaches that are utilized the most frequently. The Euclidean distance is the most straightforward approach to determining degree of similarity. Although it is simple to comprehend, it is quite sensitive to any distortions that may occur in the image. When computing Manhattan distance, rather of relying on coordinate rotation and mapping, the system's rotation of the coordinates needs to be used. The maximal-minimum distance is what is meant when people refer to the Hausdorff distance. This should not be confused with the comparable distance between two points. It's a component of fuzzy set matching, where the similarity between sets of points represents the total similarity of medical imaging features.

Modified RESNET: The initial process consisted of scaling the provided medical image up to 224 by 224 by 3 pixels in order to feed it into our CNN model. After the image had been scaled, it was sent to a deep residual CNN model so that features could be obtained by extracting from final convolutional layer. As a result of these steps, a deep feature vector with a dimension of 1*2048 was produced; this feature vector acts as a high-level representation of the intricately concealed structure that is contained inside the image that was provided as input (i.e., low-level representation). Using the Euclidean distance as a measuring stick, this extracted feature vector was put through a series of head-to-head tests against the database's labeled feature vectors. In the end, a class label was determined to be appropriate for the image that had been provided as input based on the minimal distance score. In the next sections, an explanation of the model that have been suggested is put forth, which is more comprehensive in nature.

The typical ResNet50 [18] CNN model was modified by us in our suggested framework for the classification and retrieval of medical images. One of the changes we made was to replace the model's final 7*7 average pooling layer with a 7*7*2048 convolutional layer. The use of a convolutional layer with 7*7*2048 elements is justified for the following reasons. When compared to the classification of ordinary images, the process of classifying medical images is plagued by issues associated with high levels of similarity between classes. As a result, additional features that have the potential to be useful for the classification ought to be recovered from the CNN. The original ResNet50 generates the feature map of 1*1*2048 from the prior feature map of 7*7*2048 through the use of an average pooling layer that includes one filter of 7*7, which can result in the loss of important features. Using an additional convolution layer that contained 2048 filters of 7*7*2048, our improved ResNet50 was able to overcome this issue and acquire the feature map of 1*1*2048 from the previous feature map of 7*7*2048. This helped to reduce the amount of relevant features that were lost in the process. In addition, the filter coefficients of the average pooling layer in the original ResNet50 are predetermined, however in the upgraded version of ResNet50, the optimal filter coefficients of the additional convolutional layer can be found during training. In order to demonstrate this, we conducted an experiment in which we compared the accuracy of the original ResNet50 model, which used an average pooling layer that contained one filter with a size of 7*7, to the accuracy of the revised ResNet50 model, which used an additional convolution layer that contained 2048 filters with a size of 7*7*2048. Our approach results in more accurate predictions than the original ResNet50 algorithm did.

Relevance Feedback Retrieval Method (**RFRM):** This work makes an attempt to give a reasonable representation by making use of the optical information that is latent in photographs. The process of new CBIR system is broken down and explained. The hierarchical strategy described for matching the picture template and combining it with the learning process to build detector architecture for the CBIR system was developed with the goal of making the CBIR system as efficient as possible. This strategy involves applying a bilateral filter to the photos in order to get rid of the backdrop. The edges of the objects that are present in an image can be maintained with the use of a non-linear smoothing filter like this one. After both images matching to the original image have been stored, the original image is combined with the filtered image and the image that detects shapes, and then the Canny Edge Detection algorithm is applied. Afterwards, the filtered image is used. At first, a Gaussian filter is applied to the image in order to remove any noise that may be present. After that, the image's gradients of intensity values are produced for analysis. Non maximal suppressions are applied to avoid erroneous answers. After that, we use a double threshold in order to obtain the probable edges. Hysteresis is utilized so as to track the edge points, and the edges themselves are finalized by the elimination of points that are either weak or not connected. The Euclidean distance is the method that is utilized when carrying out the calculation of proximity. This method employs metrics on vector spaces, and because of its straightforward nature, it is straightforward to put into practice. The centroid of a picture is used in the calculation of the Euclidean distance, which is then compared to the pixel that is located on the image's contour. After the shape detection stage, this image will be sent on to be analyzed for its shape descriptors, which will provide a comprehensive breakdown of the shape's essential characteristics. The shape descriptor that was created in this way is then used for all of the classes and all of the images [19]. There are a total of ten categories represented among the thousand photos that make up the Wang dataset. One hundred different pictures make up each category. The shape descriptor is checked for accuracy with the help of the SVM classifier.

Radial Associated Laguerre Moments (**RALM):** The RALMs have been suggested for use as a global descriptor in order to access the biological images. The following advantageous qualities contribute to the excellent performance of RALMs [20] in a variety of image identification applications, both in noise-free and noisy environments:
1. Orthogonal moments are represented by RALMs, and the basic functions of an orthogonal moment are also orthogonal. As a result, each RALM coefficient is able to collect distinctive and one-of-a-kind components of the image, and there is no redundant information present.
2. In both scenarios, with noise-free and noisy images, the magnitudes of the RALM coefficients are, on average, comparable to one another. As a result, RALMs would be resistant to visuals that are noisy.
3. Before and after the picture is rotated or scaled, the magnitudes of the RALM coefficients are almost the same for both versions of the image, which demonstrates that RALMs are resistant to the effects of image rotation and scaling.
4. Based on the varied order values, the basic functions of RALMs are able to extract many unique sorts of information from the image. These distinctive characteristics include the average intensity value, the variation, the information about the texture, and the information about the edges in a different

orientation. These characteristics, when taken together, offer the most accurate depiction of an image since they provide a multidimensional perspective of the image.
5. When comparable methods are used, with the exception of OFMMs, the feature values of a picture are subject to change whenever the image is rotated or resized. They are also subject to modification in the event that the photographs are noisy. The representation of those values, which represents the disparities between the center pixel and its neighbors in the spatial domain, is responsible for this change in the values of the features. These discrepancies between neighboring pixels have a propensity to shift whenever there is any modification made to the image, including rotation, scaling, and noisy instances. However, in the case of RALMs, the values of the RALMs remain the same both before and after any alteration that is made to the image due to the orthogonality property of the RALMs. Therefore, RALMs are more suited for applications that take place in the real world, such as BMIR.

III. Results and discussion

The retrieval process is considered for five classes of classification like brain scan images, bone scan images, dental scan images, lung scan images, abdomen scan images. The confusion matrix is a major evaluation matrix that should be evaluated for any medical image. Table 1 to Table 3 show various results of retrieval process with CBIR of scan images, lung scan images and abdomen scan images.

Table1: Key Performance Index evaluation of various systems - 1

Metrics / Methods	Retrieval Efficiency (%)	NARR (%)	Precision (%)	Recall (%)	F1 Score (%)	Correlation (%)	SSIM (%)	Time (sec)
TFF	75.87	51.91	73.18	73.10	73.1069	66.41	79.63	30
mRESNET	82.70	30.88	82.74	82.70	82.708	78.39	89.33	102
RFRM	94	19.58	94.05	94	94.096	92.52	94.08	28
RALM	94.60	18.98	94.61	94.60	94.59	93.25	95.16	36

Table 2: Key Performance Index evaluation of various systems - 2

Metrics / Methods	Retrieval Efficiency (%)	NARR (%)	Precision (%)	Recall (%)	F1 Score (%)	Correlation (%)	SSIM (%)	Time (sec)
TFF	74.8	51.0	72.1	72.0	72.2	65.5	78.6	27.1
mRESNET	81.6	29.9	81.6	81.6	81.8	77.5	88.3	99.1
RFRM	92.9	18.6	92.9	92.8	93.1	91.6	93.0	25.1
RALM	93.5	18.0	93.5	93.5	93.7	92.3	94.1	33.1

Table 3: Key Performance Index evaluation of various systems - 3

Metrics / Methods	Retrieval Efficiency (%)	NARR (%)	Precision (%)	Recall (%)	F1 Score (%)	Correlation (%)	SSIM (%)	Time (sec)
TFF	74.9	50.9	72.2	72.1	72.1	65.4	78.6	28.0
mRESNET	81.7	29.9	81.7	81.7	81.7	77.4	88.3	100.0
RFRM	93.0	18.6	93.0	92.9	93.0	91.5	93.1	26.0
RALM	93.6	18.0	93.6	93.6	93.6	92.3	94.1	34.0

IV. Conclusion

In this paper, it is compared a few state of art medical CBIR models and their efficiencies in terms of various quality parameters. The retrieval results of abdomen has be taken for discussion. In terms of retrieval efficiency the method RALM has produced remarkable efficiency of 93.6% and NARR of 18%. The correlation factor is also high and it is of 92.3%. The similarity index value is also high and it is of 94.1%. When come to processing speed the speed is high in RFRM model and it is of 26 seconds. The similar analysis has to be done several times on various sets of images to come to right conclusion. Based on those conclusion, a new method has to developed that must produce better results than these.

References

[1] Yousif Mohamed Y. Abdallah and Tariq Alqahtani," Research in Medical Imaging Using Image Processing Techniques" Open access peer-reviewed, 2019, DOI: 10.5772/intechopen.84360

[2] Abdallah Y. Increasing of edges recognition in cardiac scintigraphy for ischemic patients. Journal of Biomedical Engineering and Medical Imaging. 2016;

[3] Shukla, V. S., &Vala, J. A, "Survey on Image Mining, its Techniques and Application", International Journal of Computer Applications, Volume 133, (2016).

[4] Khatami, M. Babaie, H. R. Tizhoosh, A. Khosravi, T. Nguyen, and S. Nahavandi, "A sequential search-space shrinking using CNN transfer learning and a Radon projection pool for medical image retrieval," expert systems with applications, vol. 100, pp. 224–233, 2018.

[5] Radhika.K, Varadarajan.S & MMBabu, Y "Multi spectral classification using cluster ensemble technique", IJISTA, Volume:17, Issue No-1/2, pp:55-69, April-2018.

[6] A. Khatami, A. Khosravi, T. Nguyen, C. P. Lim, and S. Naha-vandi, "Medical image analysis using wavelet transform and deep belief networks," Expert Systems with Applications , vol. 86, pp. 190–198, 2017.

[7] M. Jun, Z. Senlin, and F. Zhen, "Recognition Algorithm of Fabric Based on Tamura Texture Features," Light Industry Machinery, vol. 35, no. 4, 2017.

[8] N.Koteswaramma and Y. MuraliMohanBabu "An Analysis on Medical Image Retriving Techniques", Turkish Journal of Com. and Mathematics Edu., Vol.12 No.9 April 2021, 2046 - 2053.

[9] MuraliMohan.Y, Radhika.K & Kameswari.G, "An analysis of a block matching method on single chrome images", "International Journal of Innovative Technology and Exploring Engineering", Vol.-8, No.-8, 218-220, June, 2019

[10] Nistér, D.; Stewénius, H. Scalable recognition with a vocabulary tree. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; pp. 2161–2168.

[11] Wang, J.; Liu,W.; Kumar, S.; Chang, S.-F. Learning to hash for indexing big data—A survey. Proc. IEEE 2015, 104, 34–57.

[12] Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[13] Khatami, Y. Tai, A. Khosravi et al., "A deep learning-based model for tactile understanding on haptic data percutaneous needle treatment," in International Conference on Neural Information Processing, pp. 317–325, Springer, 2017.

[14] Goodfellow, Y. Bengio, and A. Courville, Deep Learning, The MIT Press, 2016.

[15] L. Zhang, Y. Zhang, X. Gu, J. Tang, and Q. Tian, "Scalable similarity search with topology preserving hashing," IEEE Transactions on Image Processing, vol. 23, no. 7, pp. 3025-3039, 2014.

[16] Nor AsmaMohdZin, RozianiwatiYusof, Saima Anwar Lashari, Aida Mustapha, NorhalinaSenan, Rosziati Ibrahim, "Content-Based Image Retrieval in Medical Domain: A Review", IOP Conf. Series: Journal of Physics: Conf. Series 1019 (2018).

[17] Sun Xiaoming , Zhang Ning,WuHaibin, Yu Xiaoyang,WuXue, and Yu Shuang, "Medical Image Retrieval Approach by Texture Features Fusion Based on Hausdorff Distance", Hindawi-Research Article , Mathematical Problems in Engineering,2018

[18] Muhammad Owais, Muhammad Arsalan, Jiho Choi and Kang Ryoung Park, "Effective Diagnosis and Treatment through Content-Based Medical Image Retrieval (CBMIR) by Using Artificial Intelligence", MDPI-Journal of Clinical medicine, 2019, 8, 462;

[19] Ali Ahmed, "Implementing Relevance Feedback for Content-Based Medical Image Retrieval", IEEE 2020.

[20] Gaber Hassan, Khalid M. Hosny, Rushdy Mohamed Farouk And Ahmed Mansour Alzohairy, "An Efficient Retrieval System for Biomedical Images Based on Radial Associated Laguerre Moments", IEEE Access, Volume 8, 2020.